Variant Humicola Grisea CBH1.1
Goedegebuur et al.
SN# Unassigned
Docket No. GC794-2
Sheet 1 of 18

***H. grisea* CBH1.1**

Figure 1: Total Genomic Sequence (1638 nucleotides):

```
   1    ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC    50
  51    CGCCCAGCAG GCGTGCAGTC TCACCACCGA GAGGCACCCT TCCCTCTCTT   100
 101    GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC   150
 151    ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA   200
 201    CTGCTACACG GGCAACAAGT GGGATACTAG CATCTGCACT GATGCCAAGT   250
 251    CGTGCGCTCA GAACTGCTGC GTCGATGGTG CCGACTACAC CAGCACCTAT   300
 301    GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG   350
 351    CCAGCACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG   400
 401    ACAAGTATCA GAGTACGTTC TATCTTCAGC CTTCTCGCGC CTTGAATCCT   450
 451    GGCTAACGTT TACACTTCAC AGCCTTCGAG CTCCTCGGCA ACGAGTTCAC   500
 501    CTTCGATGTC GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT   550
 551    ACTTCGTCTC CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC   600
 601    AAGGCTGGTG CCAAGTACGG TACCGGCTAC TGCGATGCTC AGTGCCCCCG   650
 651    TGACATCAAG TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT   700
 701    CCACCAACGA CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT   750
 751    GAGATGGATA TCTGGGAAGC CAACAACATG GCTACTGCCT TCACTCCTCA   800
 801    CCCTTGCACC ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG   850
 851    GCACCTACAG CAACGAGCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC   900
 901    GACTTCAACT CGTACCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT   950
 951    GACCGTCGAC ACCACCAAGA AGATCACTGT CGTCACCCAG TTCCTCAAGG  1000
1001    ATGCCAACGG CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGC  1050
1051    AAGATCATCC CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAATTC  1100
1101    CATCACCCAG GACTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG  1150
1151    ACGACTTCAA CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC  1200
1201    GGCCCCATGG TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT  1250
1251    GCTCTGGCTC GACTCGACCT TCCCTGTCGA TGCCGCTGGC AAGCCCGGCG  1300
1301    CCGAGCGCGG TGCCTGCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG  1350
1351    GCCGAGGCCC CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC  1400
1401    CATCGGCTCG ACCGTTGCTG GTCTCCCCGG CGCGGGCAAC GGCGGCAACA  1450
1451    ACGGCGGCAA CCCCCCGCCC CCCACCACCA CCACCTCCTC GGCTCCGGCC  1500
1501    ACCACCACCA CCGCCAGCGC TGGCCCCAAG GCTGGCCGCT GGCAGCAGTG  1550
1551    CGGCGGCATC GGCTTCACTG GCCCGACCCA GTGCGAGGAG CCCTACACTT  1600
1601    GCACCAAGCT CAACGACTGG TACTCTCAGT GCCTGTAA               1638
```

Figure 2: Putative intron sequence deleted (GTACGTT...CAG = 413-472) Gives the cDNA sequence (1578 nucleotides):

```
   1    ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC      50
  51    CGCCCAGCAG GCGTGCAGTC TCACCACCGA GAGGCACCCT TCCCTCTCTT     100
 101    GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC     150
 151    ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA     200
 201    CTGCTACACG GGCAACAAGT GGGATACTAG CATCTGCACT GATGCCAAGT     250
 251    CGTGCGCTCA GAACTGCTGC GTCGATGGTG CCGACTACAC CAGCACCTAT     300
 301    GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG     350
 351    CCAGCACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG     400
 401    ACAAGTATCA GACCTTCGAG CTCCTCGGCA ACGAGTTCAC CTTCGATGTC     450
 451    GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT ACTTCGTCTC     500
 501    CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC AAGGCTGGTG     550
 551    CCAAGTACGG TACCGGCTAC TGCGATGCTC AGTGCCCCCG TGACATCAAG     600
 601    TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT CCACCAACGA     650
 651    CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT GAGATGGATA     700
 701    TCTGGGAAGC CAACAACATG GCTACTGCCT TCACTCCTCA CCCTTGCACC     750
 751    ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG GCACCTACAG     800
 801    CAACGAGCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC GACTTCAACT     850
 851    CGTACCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT GACCGTCGAC     900
 901    ACCACCAAGA AGATCACTGT CGTCACCCAG TTCCTCAAGG ATGCCAACGG     950
 951    CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGC AAGATCATCC    1000
1001    CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAATTC CATCACCCAG    1050
1051    GACTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG ACGACTTCAA    1100
1101    CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC GGCCCCATGG    1150
1151    TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT GCTCTGGCTC    1200
1201    GACTCGACCT TCCCTGTCGA TGCCGCTGGC AAGCCCGGCG CCGAGCGCGG    1250
1251    TGCCTGCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG GCCGAGGCCC    1300
1301    CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC CATCGGCTCG    1350
1351    ACCGTTGCTG GTCTCCCCGG CGCGGGCAAC GGCGGCAACA ACGGCGGCAA    1400
1401    CCCCCCGCCC CCCACCACCA CCACCTCCTC GGCTCCGGCC ACCACCACCA    1450
1451    CCGCCAGCGC TGGCCCCAAG GCTGGCCGCT GGCAGCAGTG CGGCGGCATC    1500
1501    GGCTTCACTG GCCCGACCCA GTGCGAGGAG CCCTACACTT GCACCAAGCT    1550
1551    CAACGACTGG TACTCTCAGT GCCTGTAA                            1578
```

Variant Humicola Grisea CBH1.1
Goedegebuur et al.
SN# Unassigned
Docket No. GC794-2
Sheet 3 of 18

Figure 3: Translation of the cDNA sequence gives the *H. grisea var thermoidea* CBH1 precursor (i.e.: with signal sequence) protein sequence (525 amino acids):

```
1       MRTAKFATLA ALVASAAAQQ ACSLTTERHP SLSWKKCTAG GQCQTVQASI        50
51      TLDSNWRWTH QVSGSTNCYT GNKWDTSICT DAKSCAQNCC VDGADYTSTY       100
101     GITTNGDSLS LKFVTKGQHS TNVGSRTYLM DGEDKYQTFE LLGNEFTFDV       150
151     DVSNIGCGLN GALYFVSMDA DGGLSRYPGN KAGAKYGTGY CDAQCPRDIK       200
201     FINGEANIEG WTGSTNDPNA GAGRYGTCCS EMDIWEANNM ATAFTPHPCT       250
251     IIGQSRCEGD SCGGTYSNER YAGVCDPDGC DFNSYRQGNK TFYGKGMTVD       300
301     TTKKITVVTQ FLKDANGDLG EIKRFYVQDG KIIPNSESTI PGVEGNSITQ       350
351     DWCDRQKVAF GDIDDFNRKG GMKQMGKALA GPMVLVMSIW DDHASNMLWL       400
401     DSTFPVDAAG KPGAERGACP TTSGVPAEVE AEAPNSNVVF SNIRFGPIGS       450
451     TVAGLPGAGN GGNNGGNPPP PTTTTSSAPA TTTTASAGPK AGRWQQCGGI       500
501     GFTGPTQCEE PYTCTKLNDW YSQCL                                  525
```

Figure 4: Mature (i.e.: expressed protein with the putative signal sequence removed) protein sequence (507 amino acids):

```
  1  QQACSLTTER HPSLSWKKCT AGGQCQTVQA SITLDSNWRW THQVSGSTNC   50
 51  YTGNKWDTSI CTDAKSCAQN CCVDGADYTS TYGITTNGDS LSLKFVTKGQ  100
101  HSTNVGSRTY LMDGEDKYQT FELLGNEFTF DVDVSNIGCG LNGALYFVSM  150
151  DADGGLSRYP GNKAGAKYGT GYCDAQCPRD IKFINGEANI EGWTGSTNDP  200
201  NAGAGRYGTC CSEMDIWEAN NMATAFTPHP CTIIGQSRCE GDSCGGTYSN  250
251  ERYAGVCDPD GCDFNSYRQG NKTFYGKGMT VDTTKKITVV TQFLKDANGD  300
301  LGEIKRFYVQ DGKIIPNSES TIPGVEGNSI TQDWCDRQKV AFGDIDDFNR  350
351  KGGMKQMGKA LAGPMVLVMS IWDDHASNML WLDSTFPVDA AGKPGAERGA  400
401  CPTTSGVPAE VEAEAPNSNV VFSNIRFGPI GSTVAGLPGA GNGGNNGGNP  450
451  PPPTTTTSSA PATTTTASAG PKAGRWQQCG GIGFTGPTQC EEPYTCTKLN  500
501  DWYSQCL                                                 507
```

```
                    1                                                                75
   CBS 225.63     (1) QQACSLTTERHPSLSWKKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVDG
D63515 mature     (1) QQACSLTTERHPSLSWKKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVDG
X17258 mature     (1) QQACSLTTERHPSLSWNKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVDG
    Consensus     (1) QQACSLTTERHPSLSWKKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVDG

                   76                                                               150
   CBS 225.63    (76) ADYTSTYGITTNGDSLSLKFVTKGQHSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM
D63515 mature    (76) ADYTSTYGITTNGDSLSLKFVTKGQYSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM
X17258 mature    (76) ADYTSTYGITTNGDSLSLKFVTKGQHSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM
    Consensus    (76) ADYTSTYGITTNGDSLSLKFVTKGQHSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM

                  151                                                               225
   CBS 225.63   (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA
D63515 mature   (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA
X17258 mature   (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA
    Consensus   (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA

                  226                                                               300
   CBS 225.63   (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKGMTVDTTKKITVVTQFLKDANGD
D63515 mature   (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKGMTVDTTKKITVVTQFLKDANGD
X17258 mature   (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKGMTVDTTKKITVVTQFLKDANGD
    Consensus   (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKGMTVDTTKKITVVTQFLKDANGD

                  301                                                               375
   CBS 225.63   (301) LGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDH
D63515 mature   (301) LGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDH
X17258 mature   (301) LGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDH
    Consensus   (301) LGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDH

                  376                                                               450
   CBS 225.63   (376) ASNMLWLDSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAGNGGNNGGNP
D63515 mature   (376) ASNMLWLDSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAGNGGNNGGNP
X17258 mature   (376) ASNMLWLDSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAGNGGNNGGNP
    Consensus   (376) ASNMLWLDSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAGNGGNNGGNP

                  451                                                          507
   CBS 225.63   (451) PPPTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL
D63515 mature   (451) PPPTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL
X17258 mature   (451) PPPTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYICTKLNDWYSQCL
    Consensus   (451) PPPTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL
```

Figure 5. Sequence alignment of two public sequences and variant *H. grisea* CBH1.1

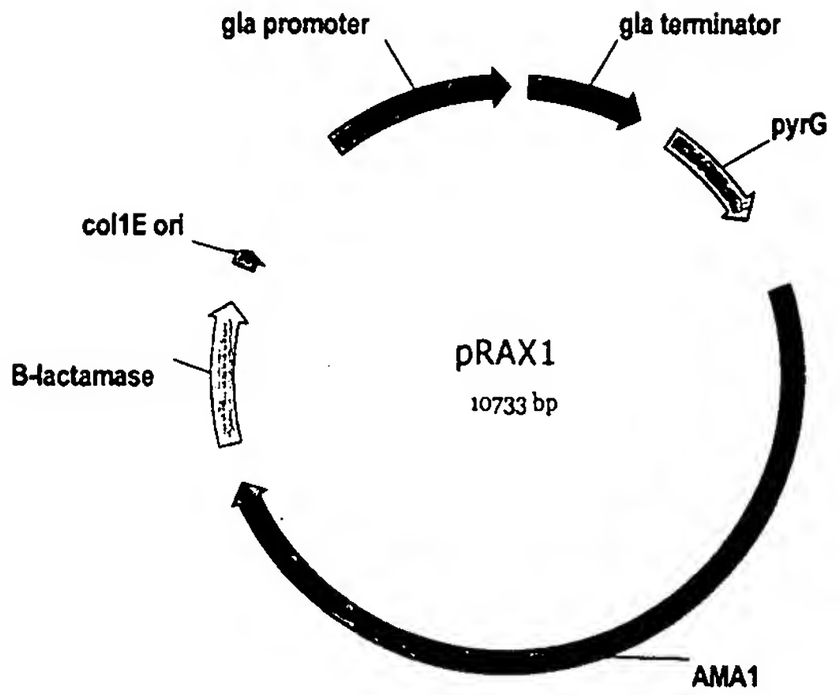


Figure 6: pRAX1

Figure 7: Destination vector pRAXdes2 for expression in *A. niger*

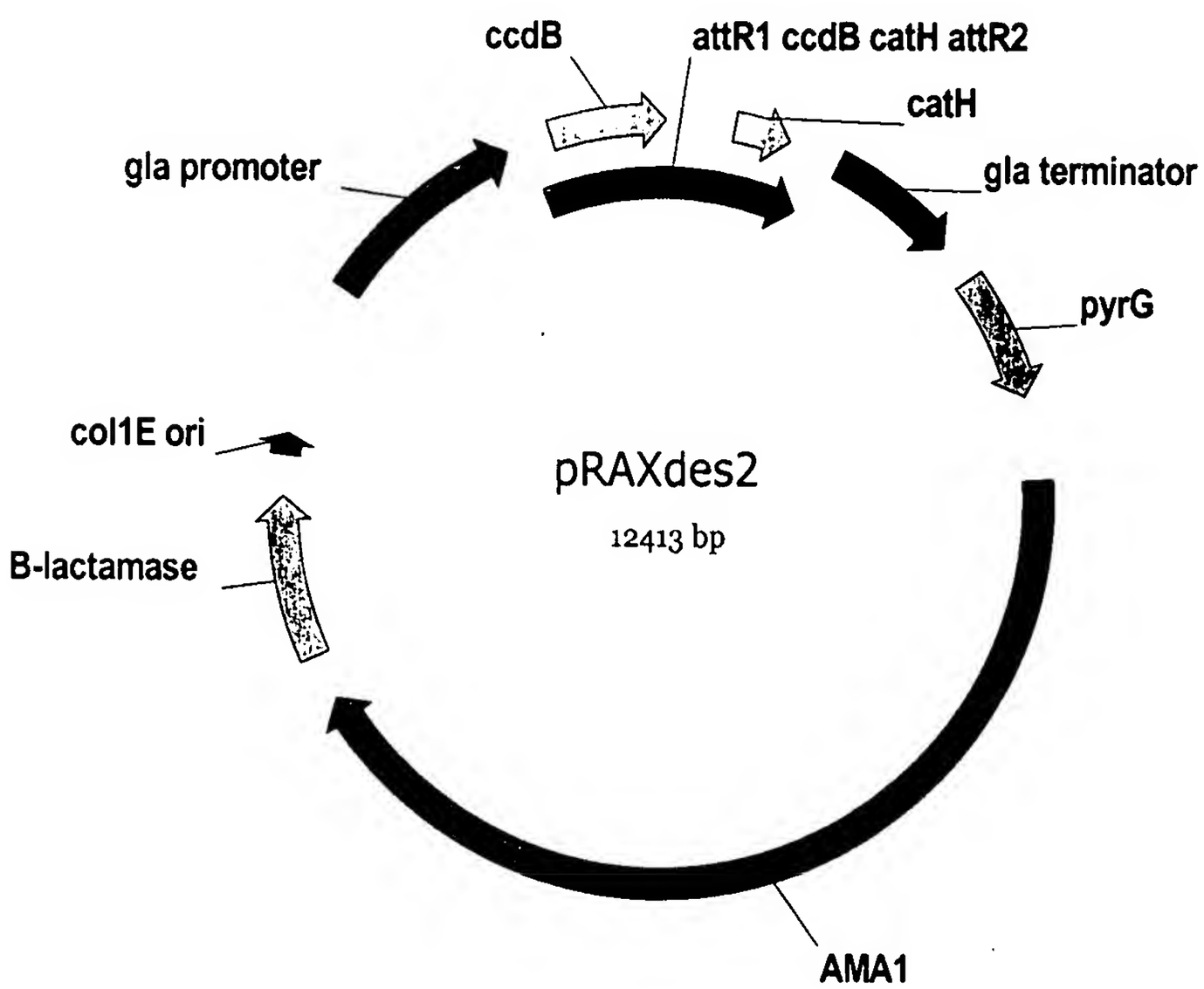


Figure 8: Replicative expression pRAXdesCBH1 vector of CBH1 genes under the control of the glucoamylase promotor.

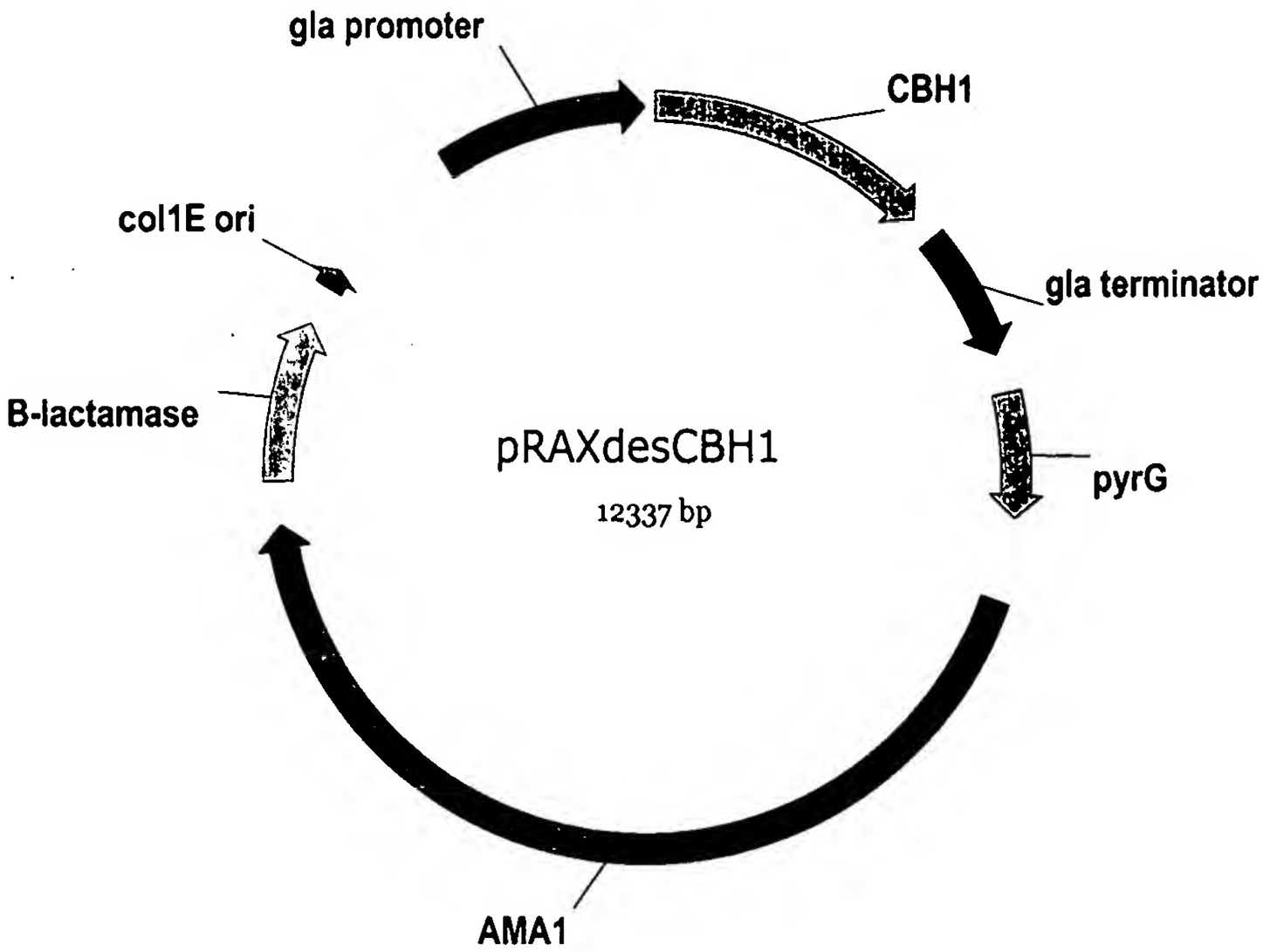


Figure 9: CBHI Conversion Assay:
38C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

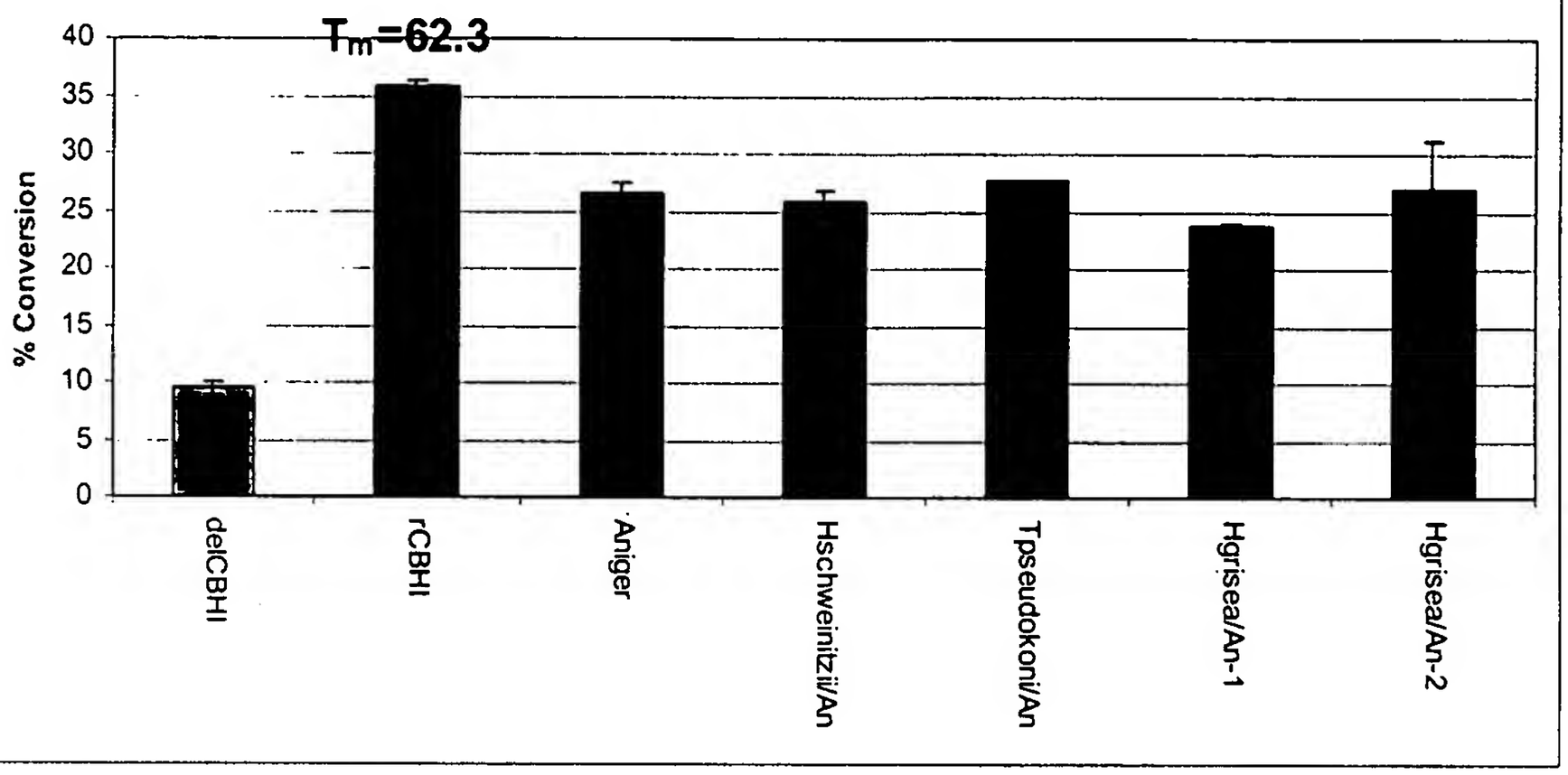


Figure 10: CBHI Conversion Assay:
65C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

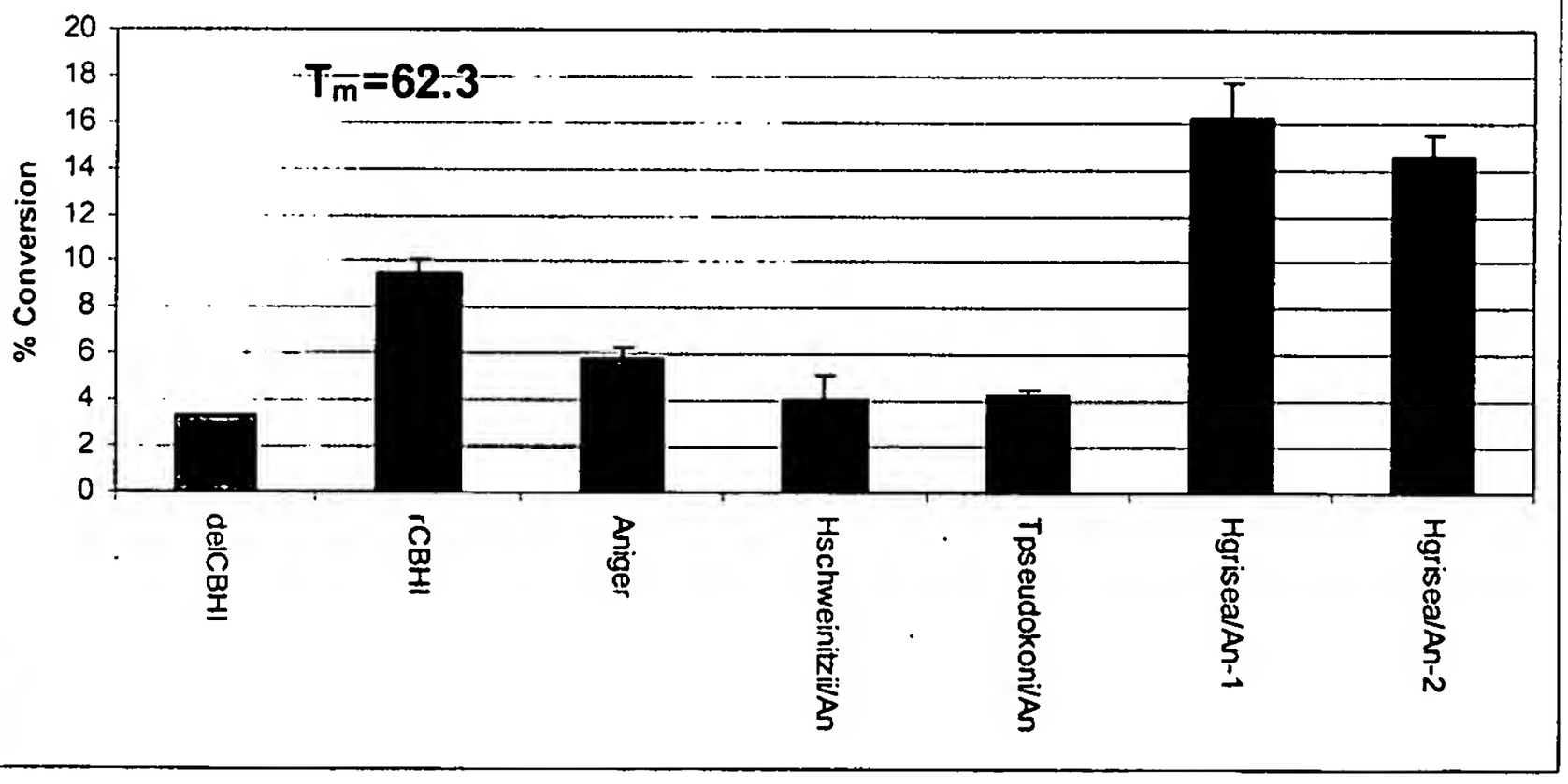


Figure 1 1: CBHI Conversion Assay:
38C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

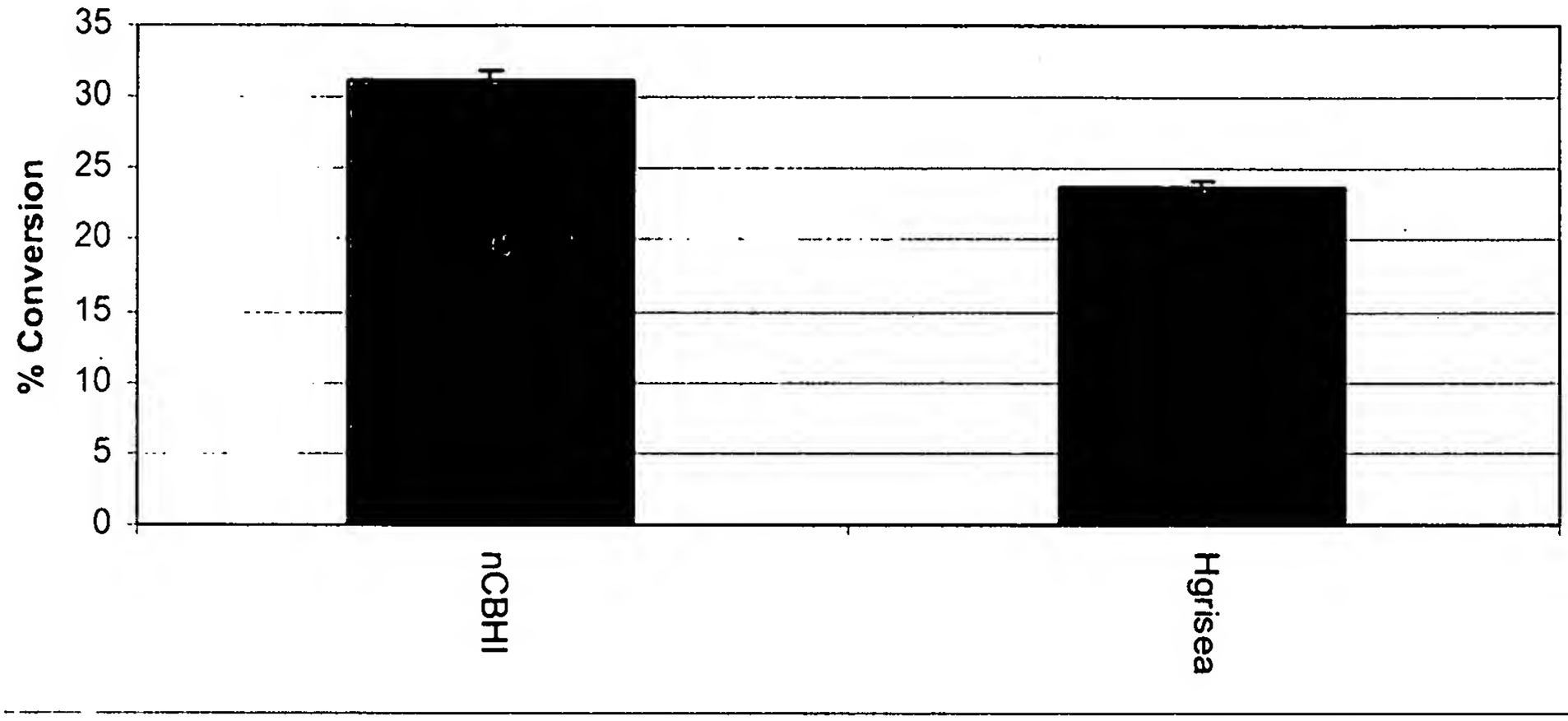


Figure 12: CBHI Conversion Assay:
65C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

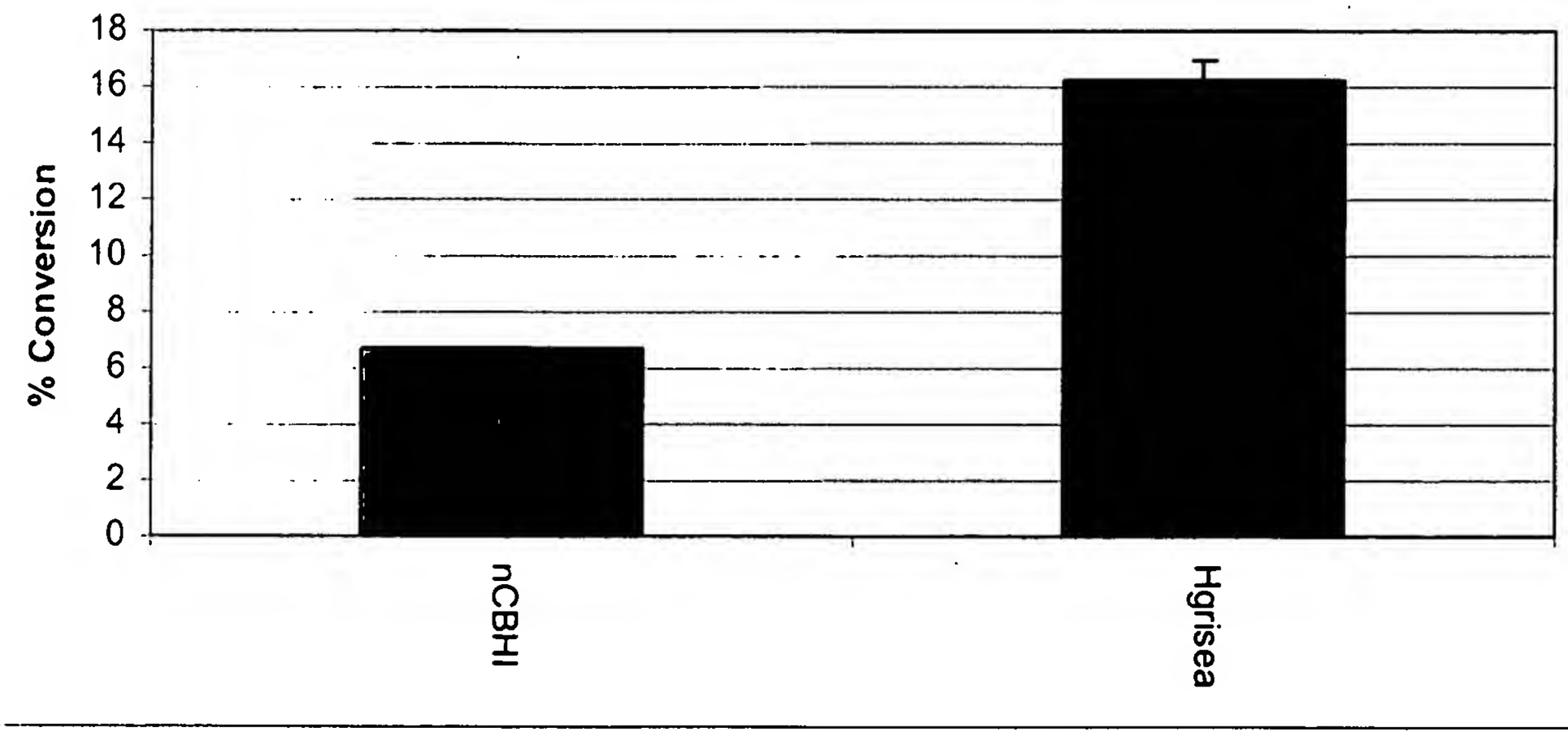


Figure 13A

3 fold rate difference b/n *H. grisea* and rCBHI

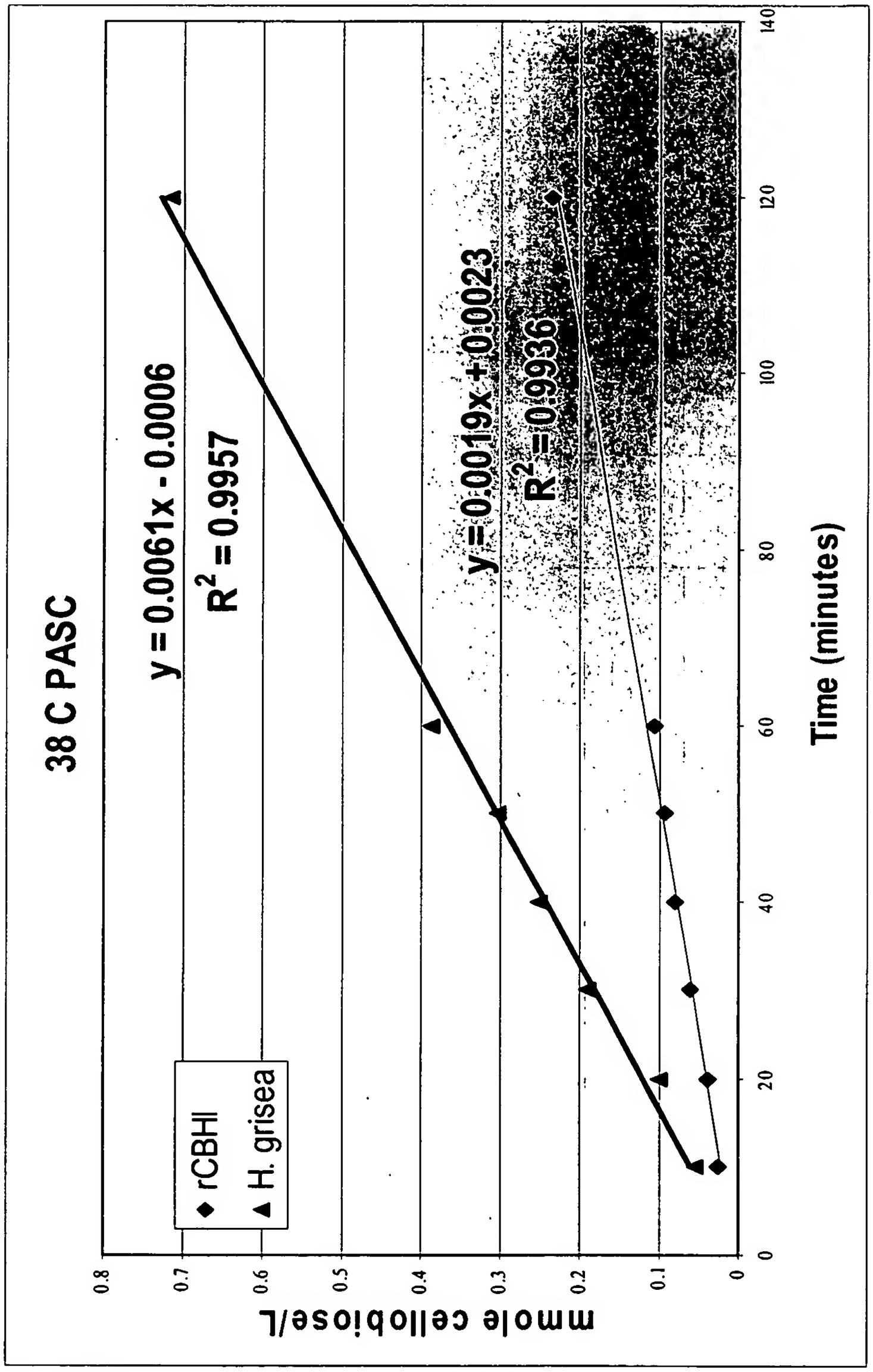


Figure 13B
4.8 fold rate difference

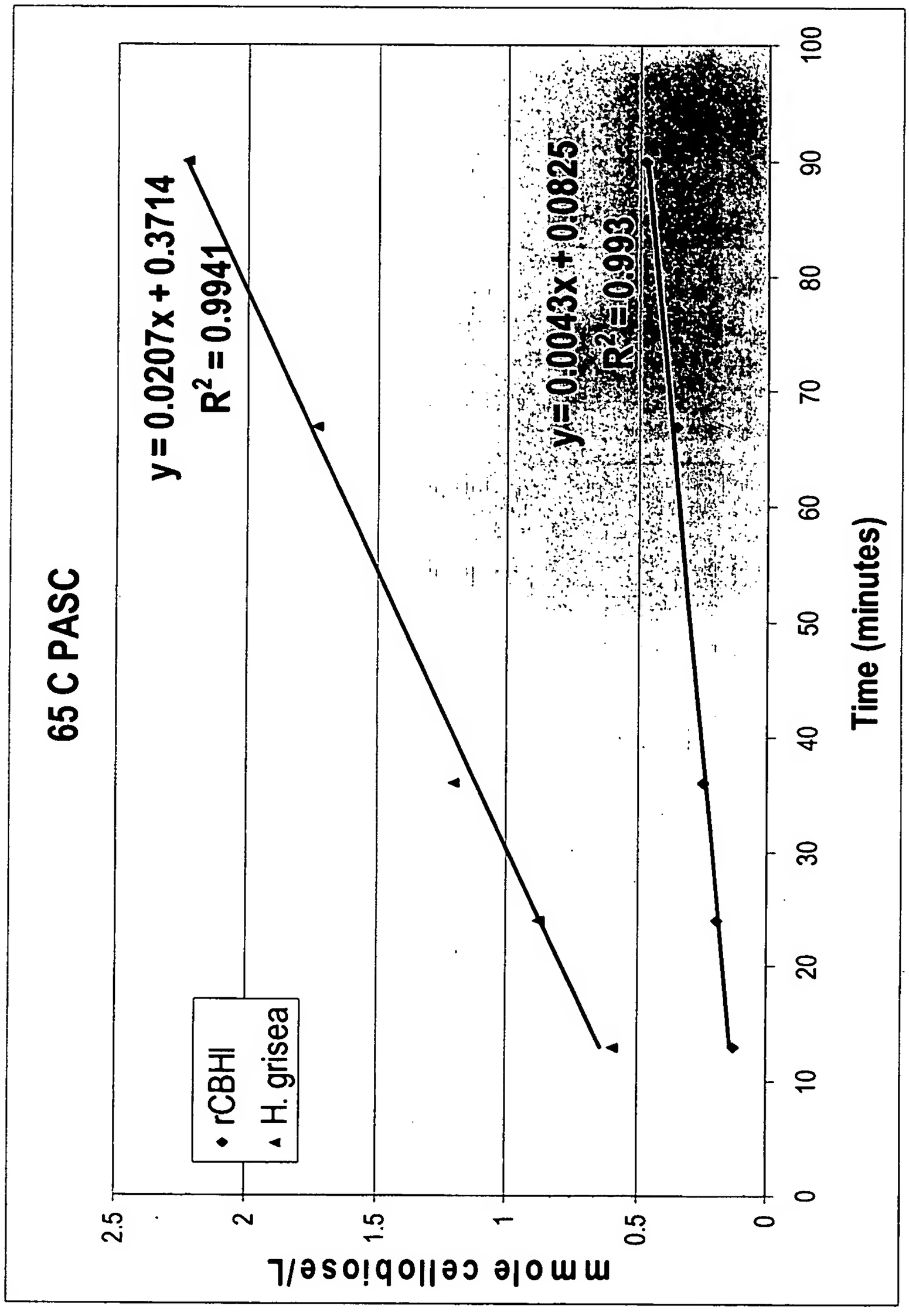
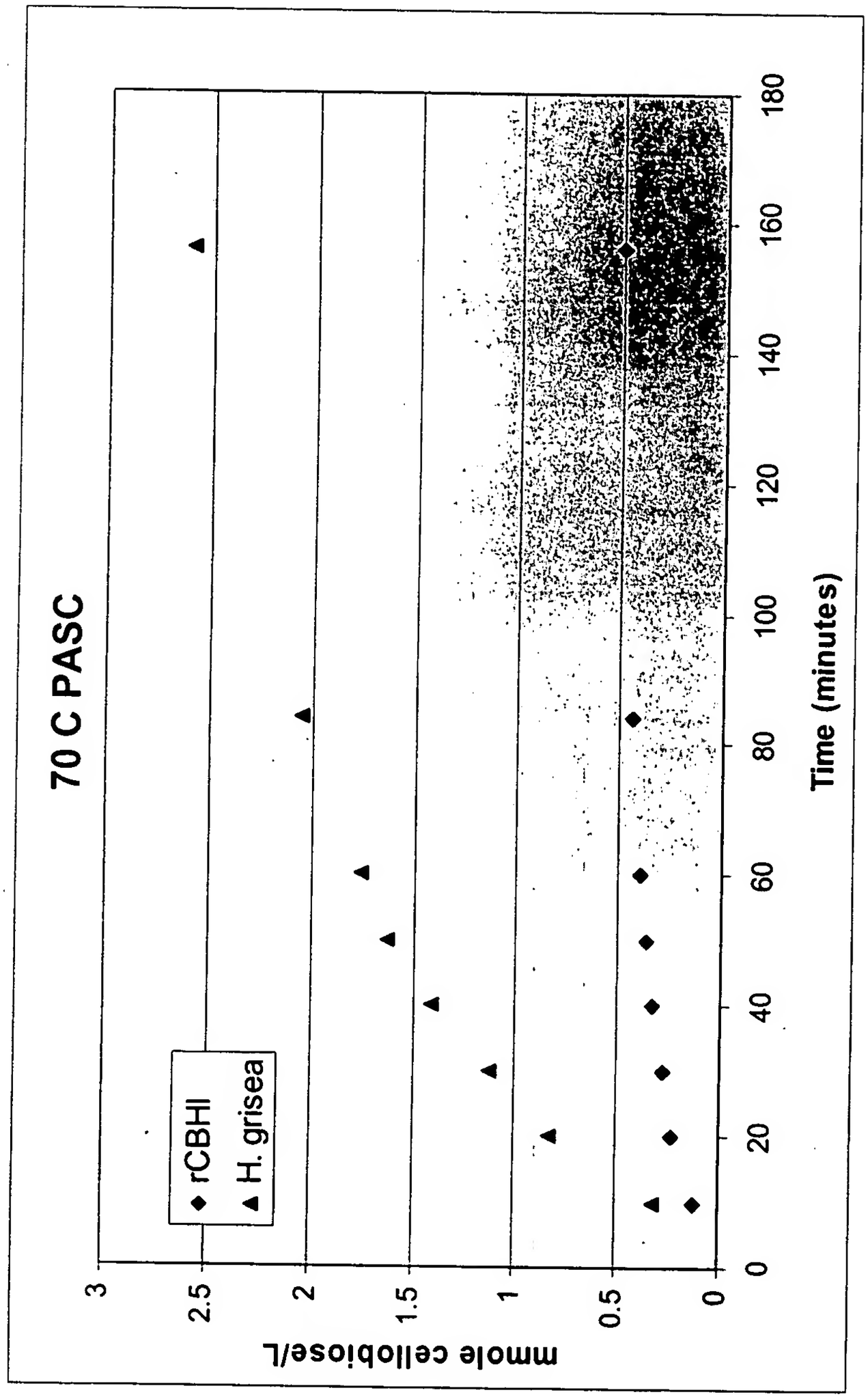


Figure 13C


70 C PASC
mmole cellobiose/L
Time (minutes)
rCBHI
H. grisea
0
0.5
1
1.5
2
2.5
3
0
20
40
60
80
100
120
140
160
180

```
   1 ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
  51 CGCCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCT TCCCTCTCCT
 101 GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
 151 ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
 201 CTGCTACACG GGCAACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
 251 CGTGCGCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
 301 GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
 351 CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
 401 ACAAGTATCA GAGTAGGTTC TATCTTCAGC CTTCTCGCGC CTTGAATCCT
 451 GGCTAACTTT TACACTTCAC AGCCTTCGAG CTCCTCGGCA ACGAGTTCAC
 501 CTTCGATGTC GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT
 551 ACTTCGTCTC CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC
 601 AAGGCTGGTG CCAAGTACGG TACCGGCTAC TGCGATGCTC AGTGCCCCCG
 651 TGACATCAAG TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT
 701 CCACCAACGA CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT
 751 GAGATGGATA TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA
 801 CCCTTGCACT ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG
 851 GCACCTACAG CAACGACCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC
 901 GACTTCAACG CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT
 951 GACCGTCGAC ACCACCAAGA AGCTCACCGT CGTCACCCAG TTCCTCAAGG
1001 ACGCCAACGG CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG
1051 AAGATCATCC CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC
1101 CATCACCCAG GATTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG
1151 ACGACTTCAA CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC
1201 GGCCCCATGG TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT
1251 GCTCTGGCTC GACTCGACCT TCCCTGTCGA TGCCGCTGGC AAGCCCGGCG
1301 CCGAGCGCGG TGCCTGCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG
1351 GCCGAGGCCC CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC
1401 CATCGGCTCG ACCGTTGCCG GCCTTCCCAG CGATGGCGGC AACAACGGCG
1451 GCAACACCAC CGTCCAGCCC CCGCCCAGCA CCACCACCAC CTCTGCCAGC
1501 AGCAGCACCA CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC
1551 CAAGGCTGGC CGCTGGCAGC AGTGCGGCGG CATCGGCTTC ACTGGCCCGA
1601 CCCAGTGCGA GGAGCCCTAC ACTTGCACCA AGCTCAACGA CTGGTACTCT
1651 CAGTGCCTGT AA
```

Figure 14A

Scytalidium thermophilium CBH1 Genomic DNA

```
   1  ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
      CGCCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCT TCCCTCTCCT
 101  GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
      ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
 201  CTGCTACACG GGCAACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
      CGTGCGCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
 301  GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
      CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
 401  ACAAGTATCA GACCTTCGAG CTCCTCGGCA ACGAGTTCAC CTTCGATGTC
      GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT ACTTCGTCTC
 501  CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC AAGGCTGGTG
      CCAAGTACGG TACCGGCTAC TGCGATGCTC AGTGCCCCCG TGACATCAAG
 601  TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT CCACCAACGA
      CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT GAGATGGATA
 701  TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA CCCTTGCACT
      ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG GCACCTACAG
 801  CAACGACCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC GACTTCAACG
      CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT GACCGTCGAC
 901  ACCACCAAGA AGCTCACCGT CGTCACCCAG TTCCTCAAGG ACGCCAACGG
      CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG AAGATCATCC
1001  CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC CATCACCCAG
      GATTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG ACGACTTCAA
1101  CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC GGCCCCATGG
      TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT GCTCTGGCTC
1201  GACTCGACCT TCCCTGTCGA TGCCGCTGGC AAGCCCGGCG CCGAGCGCGG
      TGCCTGCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG GCCGAGGCCC
1301  CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC CATCGGCTCG
      ACCGTTGCCG GCCTTCCCAG CGATGGCGGC AACAACGGCG GCAACACCAC
1401  CGTCCAGCCC CCGCCCAGCA CCACCACCAC CTCTGCCAGC AGCAGCACCA
      CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC CAAGGCTGGC
1501  CGCTGGCAGC AGTGCGGCGG CATCGGCTTC ACTGGCCCGA CCCAGTGCGA
      GGAGCCCTAC ACTTGCACCA AGCTCAACGA CTGGTACTCT CAGTGCCTGT
1601  AA
```

Figure 14B
Scytalidium thermophilium CBH1 cDNA

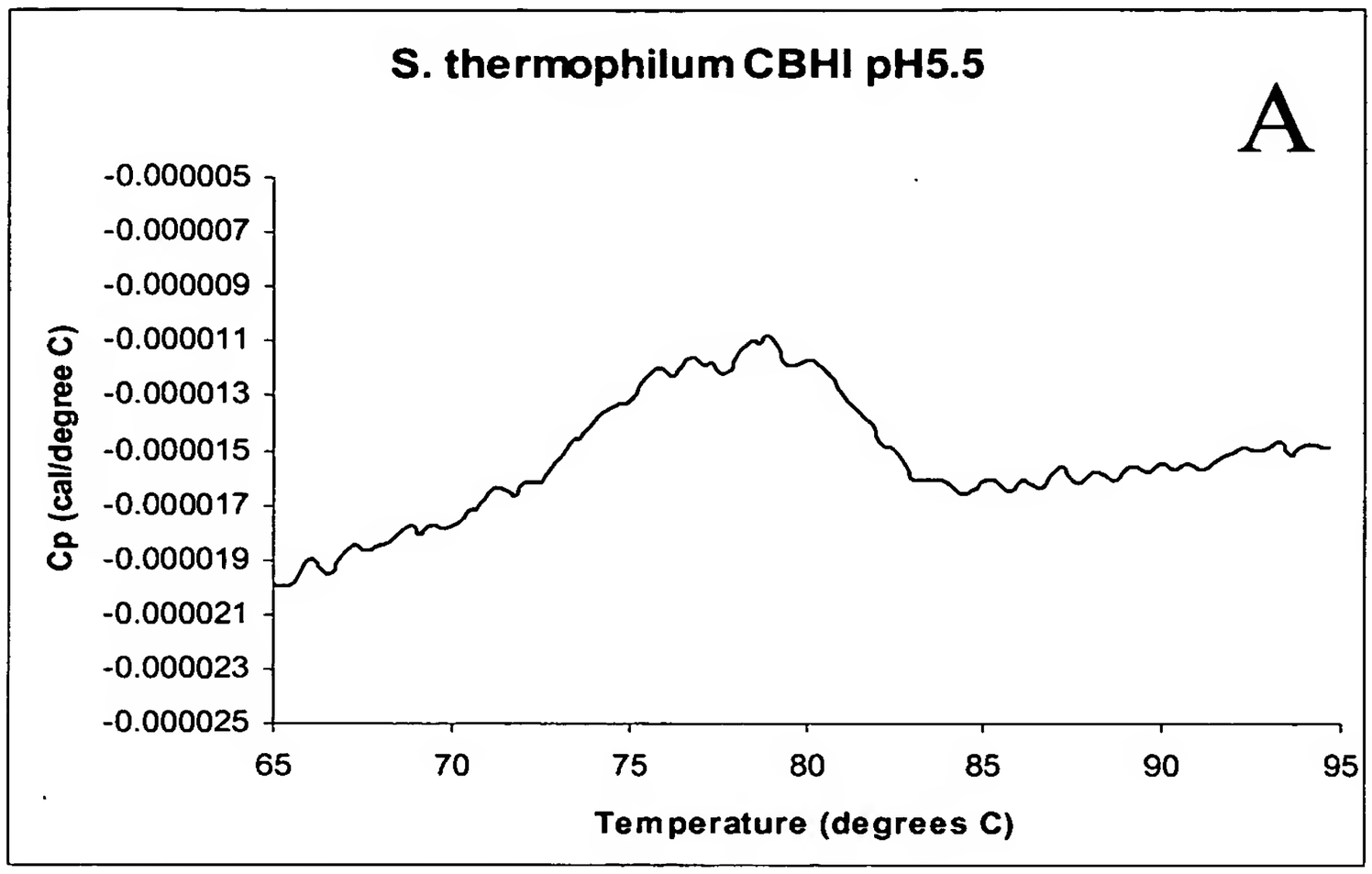
```
1    MRTAKFATLAALVASAAAQQACSLTTERHPSLSWKKCTAGGQCQTVQASI   50

51   TLDSNWRWTHQVSGSTNCYTGNEWDSSICTDAKSCAQNCCVDGADYTSTY  100

101  GITTNGDSLSLKFVTKGQYSTNVGSRTYLMDGEDKYQTFELLGNEFTFDV  150

151  DVSNIGCGLNGALYFVSMDADGGLSRYPGNKAGAKYGTGYCDAQCPRDIK  200

201  FINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATAFTPHPCT  250

251  IIGQSRCEGDSCGGTYSNDRYAGVCDPDGCDFNAYRQGNKTFYGKGMTVD  300

301  TTKKLTVVTQFLKDANGDLGEIKRFYVQDGKIIPNSESTIPGVEGNSITQ  350

351  DWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDHASNMLWL  400

401  DSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGS  450

451  TVAGLPSDGGNNGGNTTVQPPPSTTTTSASSSTTSAPATTTTASAGPKAG  500

501  RWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL-                   534
```

Figure 14C

Scytalidium thermophilum CBHI, including signal sequence

Figure 15: Alignment of the mature protein sequences for *Hypocrea jecorina* CBHI, *Humicola grisea var. thermoidea* CBH1.1, and *Scytalidium thermophilum* CBH

```
                                        1                                                                  75
Genencor Hypocrea jecorina Cel7A    (1) QSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVIDANWRWTHATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDG
Humicola grisea CBH1.1              (1) QQACSLTTERHPSLSWNKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVDG
Scytalidium thermophilum 69         (1) QQACSLTTERHPSLSWKKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNEWDSSICTDAKSCAQNCCVDG
                          Consensus (1) QQACSLTTERHPSLSWKKCTAGGQCQTVQASITLDSNWRWTHQVSGSTNCYTGNKWDSSICTDAKSCAQNCCVDG

                                       76                                                                 150
Genencor Hypocrea jecorina Cel7A   (76) AAYASTYGVTTSGNSLSIGFVTQSAQKN-VGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLPCGLNGALYFVSM
Humicola grisea CBH1.1             (76) ADYTSTYGITTNGDSLSLKFVTKGQHSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM
Scytalidium thermophilum 69        (76) ADYTSTYGITTNGDSLSLKFVTKGQYSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM
                          Consensus (76) ADYTSTYGITTNGDSLSLKFVTKGQHSTNVGSRTYLMDGEDKYQTFELLGNEFTFDVDVSNIGCGLNGALYFVSM

                                      151                                                                 225
Genencor Hypocrea jecorina Cel7A  (150) DADGGVSKYPTNTAGAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSNNANTGIGGHGSCCSEMDIWEANSISEA
Humicola grisea CBH1.1            (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA
Scytalidium thermophilum 69       (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA
                          Consensus (151) DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIKFINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATA

                                      226                                                                 300
Genencor Hypocrea jecorina Cel7A  (225) LTPHPCTTVGQEICEGDGCGGTYSDNRYGGTCDPDGCDWNPYRLGNTSFYGPGSSFTLDTTKKLTVVTQFETSG-
Humicola grisea CBH1.1            (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKG--MTVDTTKKITVVTQFLKDAN
Scytalidium thermophilum 69       (226) FTPHPCTIIGQSRCEGDSCGGTYSNDRYAGVCDPDGCDFNAYRQGNKTFYGKG--MTVDTTKKLTVVTQFLKDAN
                          Consensus (226) FTPHPCTIIGQSRCEGDSCGGTYSNERYAGVCDPDGCDFNSYRQGNKTFYGKG  MTVDTTKKITVVTQFLKDAN

                                      301                                                                 375
Genencor Hypocrea jecorina Cel7A  (299) ----AINRYYVQNGVTFQQPNAELGSYSGNELNDDYCTAEEAEFGGSS-FSDKGGLTQFKKATSGGMVLVMSLWD
Humicola grisea CBH1.1            (299) GDLGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWD
Scytalidium thermophilum 69       (299) GDLGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWD
                          Consensus (301) GDLGEIKRFYVQDGKIIPNSESTIPGVEGNSITQDWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWD

                                      376                                                                 450
Genencor Hypocrea jecorina Cel7A  (369) DYYANMLWLDSTYPTNETSSTPGAVRGSCSTSSGVPAQVESQSPNAKVTFSNIKFGPIGSTGNPSGGNP------
Humicola grisea CBH1.1            (374) DHASNMLWLDSTFPVDAAG-KPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAG--NGGN
Scytalidium thermophilum 69       (374) DHASNMLWLDSTFPVDAAG-KPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPSDGGNNGGN
                          Consensus (376) DHASNMLWLDSTFPVDAAG KPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGSTVAGLPGAG  NGGN

                                      451                                                                 518
Genencor Hypocrea jecorina Cel7A  (438) PGGNPPG--------TTTTRRPATTTGSSPGPTQSHYGQCGGIGYSGPTVCASGTTCQVLNPYYSQCL
Humicola grisea CBH1.1            (446) NGGNPPP------PTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYICTKLNDWYSQCL
Scytalidium thermophilum 69       (448) TTVQPPPSTTTTSASSSTTSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL
                          Consensus (451) NGGNPPP      PTTTTSSAPATTTTASAGPKAGRWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL
```


S. thermophilum CBHI pH5.5
A
Cp (cal/degree C)
-0.000005
-0.000007
-0.000009
-0.000011
-0.000013
-0.000015
-0.000017
-0.000019
-0.000021
-0.000023
-0.000025
65
70
75
80
85
90
95
Temperature (degrees C)

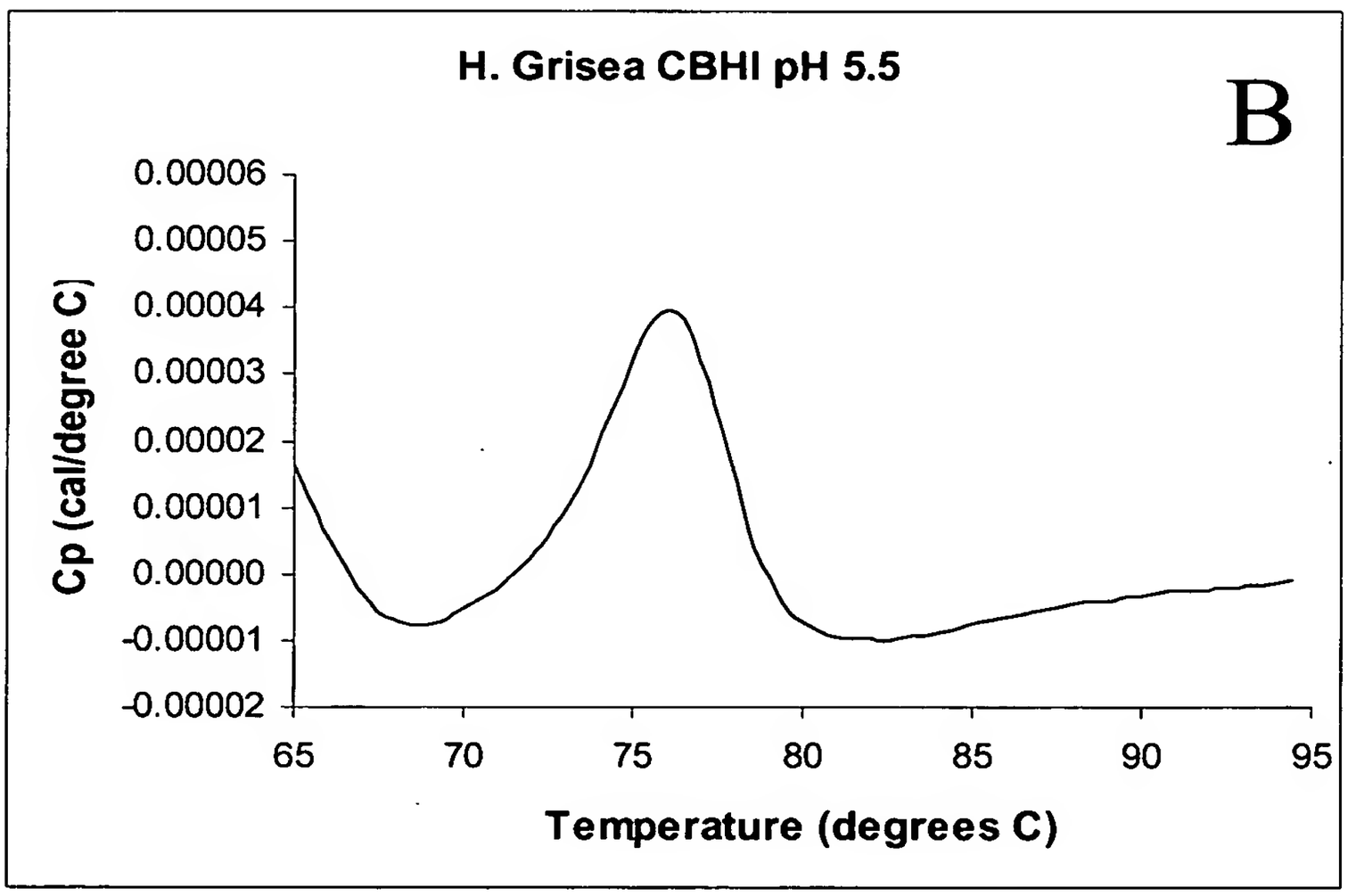

H. Grisea CBHI pH 5.5
B
Cp (cal/degree C)
0.00006
0.00005
0.00004
0.00003
0.00002
0.00001
0.00000
-0.00001
-0.00002
65
70
75
80
85
90
95
Temperature (degrees C)

Figure 16